

TPLGPASSLPQSFLCLKLEQVRKIQGDGAALQEKLCATYKLCHPPELVLLGHSLGIPWA
PLSSCPSQLQLAGCLSQLHSGFLFYQGLLQALEGISPELGPTLDTLQLDVAADFATTIW
QQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSRYRVLRLHAQP

hGCSF wild type

ATGACTCCATTAGGTCCAGCTTCCTCTCTGCCCCAAAGCTTCCTGCTGAAATGCCCTGGAACAGGTTCGTAAATCCAGGGTGATGG
TGCTGCTCTGCAGGAAAAACTGTGCGCTACCTACAAACTGTGCCATCCGGAAGAACTGGTTCTGCTGGGTCACCTCCCTGGGTATCC
CGTGGGCGCCGCTGAGCTCCTGCCCCGAGCCAGGCTCTGCAGCTGGTGGTCCCAATTGCACAGCGGCCCTTTTCCCTGTAC
CAGGGTCTGCTGCAAGCTCTGGAAGGTATCTCCCGGAACCTGGGTCCGACACTCTGCAGCTGGACGTCGCTGACTTCGC
TACCACCATCTGGCAGCAGATGGAAGAACTGGGTATGGCTCCGGCTCTGCAGCCGACCCAGGGTGCTATGCCGGCTTTCGCTTCCG
CTTTCAGCGTCGCGCAGGTGGCGTTCTGGTTGCTAGCCACCTGCAGAGCTTCCCTGGAAGTTTCCCTACCGTGTCTGCGTCACCTG
GCTCAGCCGCTGA

FIG. 1

core3	17 95	21 99	24 103	28 106	31 110	35 113	41 114	47 117	54 140	56 151	75 152	78 153	82 154	85 157	88 160	89 161	92 168
core4	17 150	21 151	24 152	28 153	31 154	75 157	78 160	82 161	85 168	89	103	106	110	113	114	117	149
core4v	17 152	21 153	24 154	28 157	31 160	75 161	78 167	82 168	85	89	103	106	110	113	114	117	151
bndry4_2	14 120	20 145	27 146	32 147	34 148	38 155	77 156	79 164	84 170	91	99	102	107	109	116		
bndry4_core4	14 120	20 145	27 146	32 147	34 148	38 155	77 156	79 164	84 170	91	99	102	107	109	116		
bndry4_AD	14	20	27	32	34	38	145	146	147	148	155	156	164	170			
bndry4_AD_core4	14	20	27	32	34	38	145	146	147	148	155	156	164	170			

FIG._2

G-CSF Designs - Optimal Sequences Selected by PDA*

	1	10	20	30	40	50	60
hGCSFwt	MTPLGPASSLPQSFLLKCLEQVRKIQDGAALQEKLCATYKLCHPPELVLLGHSGLGPWAP						
bndry4_2		I	L	E	I	I	K
bndry4_core4		I	L	EA	L	E	H
bndry4_AD		I	L	E	I	E	H
bndry4_AD_core4		I	L	EA	L	E	H
core4		L	L	A			
core4_V167A		L	I	A			
core3		L	L	A	I		
sm0		A	A	A			
fm2		A	A	A			
fm3		L	A	A			
fm4		L	L	A			
fm7		L	L	A			

	70	80	90	100	110	120
hGCSFwt	LSSCPSQALQLAGCLSQLHSGFLFYQGLLQALEGISPELGPTLDTLQLDVA					
bndry4_2		L	K	V	I	I
bndry4_core4		FL	K	KV	I	I
bndry4_AD		F	F	I	I	I
bndry4_AD_core4		F	F	V	I	L
core4		F	F	V	I	L
core4_V167A		F	F		I	L
core3		F	F	F	L	
sm0		F	F	F		
fm2		F				
fm3		F				
fm4		F				
fm7		F				

FIG._3A

	130	140	150	160	170
hGCSFwt	MEELGMAPALQPTQ	GAMPAPAFASAFQ	RAGGVLVASHLQ	SFLEVSYRVLRH	LAQP
bndry4_2			KED	IL	A
bndry4_core4			KED	I I IL	A F
bndry4_AD			KET	IL	A
bndry4_AD_core4			KED	I I IL	A F
core4			I I		F
core4_V167A			I I	I WF	AF
core3			I I		F
sm0					
fm2			I I		F
fm3					
fm4			I I		F
fm7			I I		F

*Sequences shown below dotted lines were not obtained from PDA calculations but were derived by reverting some core4 or core3 mutant positions to wild type. Core4 mutant positions are indicated in bold. The sequence selected for Core4_V167A is not the ground state; Monte Carlo analysis shows the ground state with Phe instead of Trp for position 160, and Leu instead of Phe for position 161 (see Table 4).

FIG._3B

FIG. 4

Core4 - Monte Carlo Analysis - Ground State and Allowed Amino Acids
and Their Number of Occurrences (For the Top 1000 Sequences)

hG-CSF	Position	Ground State		
CYS	17	LEU 736	ILE 229	
VAL	21	VAL 687	ILE 287	
ILE	24	VAL 38	ILE 961	
GLY	28	ALA 747	LEU 172	
LEU	31	VAL 251	LEU 707	
LEU	75*	LEU 999		
LEU	78	PHE 974		
LEU	82*	LEU 974		
TYR	85	PHE 847	TYR 140	
LEU	89	LEU 628	PHE 321	
LEU	103	VAL 351	LEU 264	ILE 313
LEU	106*	LEU 940		
VAL	110	VAL 415	LEU 143	ILE 441
PHE	113	LEU 999		
ALA	114*	ALA 999		
ILE	117*	ILE 956		
GLY	149*	GLY 999		
GLY	150*	GLY 999		
VAL	151	ILE 999		
LEU	152*	LEU 999		
VAL	153	VAL 411	ILE 588	
ALA	154*	ALA 999		
LEU	157	LEU 805	ILE 187	
PHE	160	PHE 565	TRP 434	
LEU	161	LEU 838	PHE 161	
LEU	168	PHE 999		

*position where Monte Carlo didn't find an alternative and where the top amino acid is the wild type

Table 5. Core3 - Monte Carlo Analysis (Ground State and Allowed Amino Acids and Their Number of Occurrences (For the Top 1000 Sequences)

hG-CSF	Position	Ground State										
CYS	17	LEU 585	VAL 35	ILE 379								
VAL	21	VAL 551	ALA 15	ILE 291	PHE 141	TYR 1						
ILE	24	ILE 657	ALA 31	VAL 303	LEU 8							
GLY	28	ALA 928	LEU 71									
LEU	31	LEU 888	VAL 111									
LYS	35	ILE 785	VAL 214									
LEU	41*	LEU 999										
LEU	47*	LEU 999										
LEU	54*	LEU 999										
ILE	56*	ILE 999										
LEU	75*	LEU 999										
LEU	78	PHE 692	ALA 10	VAL 149	LEU 12	ILE 75	TYR 61					
LEU	82	LEU 851	ALA 12	PHE 136								
TYR	85	PHE 636	TRP 363									
LEU	88*	LEU 999										
LEU	89	PHE 674	LEU 214	TRP 111								
LEU	92	PHE 999										
ILE	95*	ILE 999										
LEU	99*	LEU 999										
LEU	103	LEU 888	ILE 111									
LEU	106	LEU 893	VAL 106									
VAL	110	VAL 400	ALA 14	LEU 294	ILE 291							
PHE	113	LEU 954	ALA 1	PHE 44								
ALA	114*	ALA 999										
ILE	117	ILE 790	ALA 15	VAL 168	LEU 5	PHE 20	TRP 1					

FIG.-6A

Table 5. Core3 - Monte Carlo Analysis (Ground State and Allowed Amino Acids and Their Number of Occurrences (For the Top 1000 Sequences))

hG-CSF	Position	Ground State
PHE	140*	PHE 999
VAL	151	ILE 999
LEU	152*	LEU 999
VAL	153	ILE 999
ALA	154*	ALA 999
LEU	157	LEU 694
PHE	160	PHE 574
LEU	161	LEU 784
LEU	168	PHE 999

ALA 22 VAL 179 ILE 104
TRP 425
ALA 6 VAL 55 PHE 154

FIG. 6B

*position where Monte Carlo didn't find an alternative and where the top amino acid is the wild type

Table 6. Bndry4_2 - Monte Carlo Analysis (Ground State and Allowed Amino Acids and Their Number of Occurrences (For the Top 1000 Sequences))

hG-CSF	Position	Ground State
LEU	14	ILE 998
GLN	20	LEU 999
ASP	27	GLU 999
GLN	32	ILE 999
LYS	34	LYS 717
THR	38	VAL 409
GLN	77*	GLN 999
HIS	79	LEU 999
LEU	84*	LEU 999
ALA	91	LYS 999
LEU	99	VAL 759
THR	102	LEU 562
GLN	107	ILE 993
VAL	109	GLU 525
THR	116	ILE 749
GLN	120	LEU 999
GLN	145	GLN 650
ARG	146	LYS 891
ARG	147	GLU 999
ALA	148	THR 401
SER	155	ILE 999
HIS	156	LEU 999
SER	164	ALA 999
HIS	170	HSP 380
		ILE 209 GLU 73
		ILE 188 GLU 237 LYS 154
		LEU 193
		ILE 404
		VAL 474
		LEU 198 LYS 52
		GLU 349
		GLN 108
		ALA 268 ASP 330
		LEU 111 GLU 248 GLN 227

*position where Monte Carlo didn't find an alternative and where the top amino acid is the wild type

FIG. 7

9

position where Monte Carlo didn't find an alternative and where the top amino acid is the wild type

FIG.--9

Table 8. Bndry4_AD - Monte Carlo Analysis (Ground State and Allowed Amino Acids and Their Number of Occurrences (For the Top 1000 Sequences))

Position	Ground State	
14	ILE 887	LEU 112
20	LEU 999	
27	GLU 984	
32	ILE 931	
34	GLU 357	ILE 68
38	VAL 287	ILE 133
145	GLN 605	GLN 223
146	LYS 786	HSP 225
147	GLU 962	LYS 277
148	THR 373	GLU 217
155	ILE 976	LYS 123
156	LEU 994	
164	ALA 999	
170	HSP 304	ALA 305
		ASP 321
		LEU 136
		GLU 230
		GLN 209
		LYS 62

*position where Monte Carlo didn't find an alternative and where the top amino acid is the wild type

*position where Monte Carlo didn't find an alternative and where the top amino acid is the wild type

*position where Monte Carlo didn't find an alternative and where the top amino acid is the wild type

Core3

ATGACTCCATTAGGTCCAGCTTCCCTCTCTGCGCGAAAGCTTCCCTGTGAACCTGCTGGAACAGGTTCTGTAAATCCAGGGTGATGC
 AGCTGCTCTGCAGGAAATAATCTGCGCTACCTACAAACTGTGCCATCCGGAAGAACTGGTTCTGCTGGGTCACTCCCTGGGTATCC
 CGTGGCGCGCTGAGCTCCTGCGCGAGCCAGGCTCTGCAGCTGGCTGGTTGCCCTGTCCCAATTCCACAGCGGCCCTTTTCCTGTTC
 CAGGGTCTGTTCCAGGCTTTCGAAGGTATCTCCCGGAACCTGGGTCCGACCTCTGCAGCTGGACGTGCGTGACCTGGC
 TACCACCATCTGGCAGCAGATGGAAGAACTGGGTATGGCTCCGGCTCTGCAGCCGACCCAGGGTCTATGCCGGCTTTCGGCTTCCG
 CTTTCCAGCGTCGCGCAGGTGGCATCCTGATCGCTAGCCACCTGCAGAGCTTCCCTGGAAGTTTCCCTACCGTGTTTTCCGTCACCTG
 GCTCAGCCCGTGA

FIG._ 11A

Core4

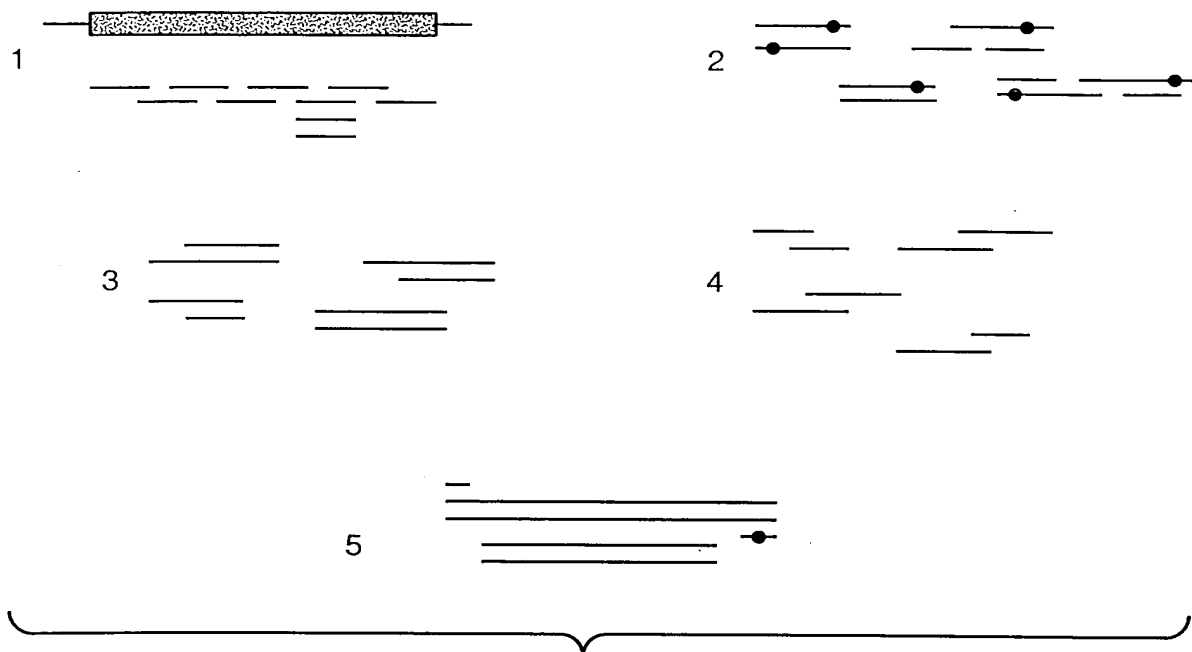
ATGACTCCATTAGGTCCAGCTTCCCTCTCTGCGCGAAAGCTTCCCTGTGAACCTGCTGGAACAGGTTCTGTAAATCCAGGGTGATGC
 AGCTGCTCTGCAGGAAATAATCTGCGCTACCTACAAACTGTGCCATCCGGAAGAACTGGTTCTGCTGGGTCACTCCCTGGGTATCC
 CGTGGCGCGCTGAGCTCCTGCGCGAGCCAGGCTCTGCAGCTGGCTGGTTGCCCTGTCCCAATTCCACAGCGGCCCTTTTCCTGTTC
 CAGGGTCTGCTGCAAGCTCTGGAAGGTATCTCCCGGAACCTGGGTCCGACCGTTGACACTCTGCAGCTGGACATCGCTGACCTGGC
 TACCACCATCTGGCAGCAGATGGAAGAACTGGGTATGGCTCCGGCTCTGCAGCCGACCCAGGGTCTATGCCGGCTTTCGGCTTCCG
 CTTTCCAGCGTCGCGCAGGTGGCATCCTGATCGCTAGCCACCTGCAGAGCTTCCCTGGAAGTTTCCCTACCGTGTTTTCCGTCACCTG
 GCTCAGCCCGTGA

FIG._ 11B

Core4v

ATGACTCCATTAGGTCCAGCTTCCCTCTCTGCGCGAAAGCTTCCCTGTGAACCTGCTGGAACAGATCCGTAAATCCAGGGTGATGC
 AGCTGCTCTGCAGGAAATAATCTGCGCTACCTACAAACTGTGCCATCCGGAAGAACTGGTTCTGCTGGGTCACTCCCTGGGTATCC
 CGTGGCGCGCTGAGCTCCTGCGCGAGCCAGGCTCTGCAGCTGGCTGGTTGCCCTGTCCCAATTCCACAGCGGCCCTTTTCCTGTTC
 CAGGGTCTGCTGCAAGCTCTGGAAGGTATCTCCCGGAACCTGGGTCCGACCTCTGCAGCTGGACATCGCTGACCTGGC
 TACCACCATCTGGCAGCAGATGGAAGAACTGGGTATGGCTCCGGCTCTGCAGCCGACCCAGGGTCTATGCCGGCTTTCGGCTTCCG
 CTTTCCAGCGTCGCGCAGGTGGCATCCTGATCGCTAGCCACATCCAGAGCTGGTTCTGGAAGTTTCCCTACCGTGTTTTCCGTCACCTG
 GCTCAGCCCGTGA

FIG._ 11C

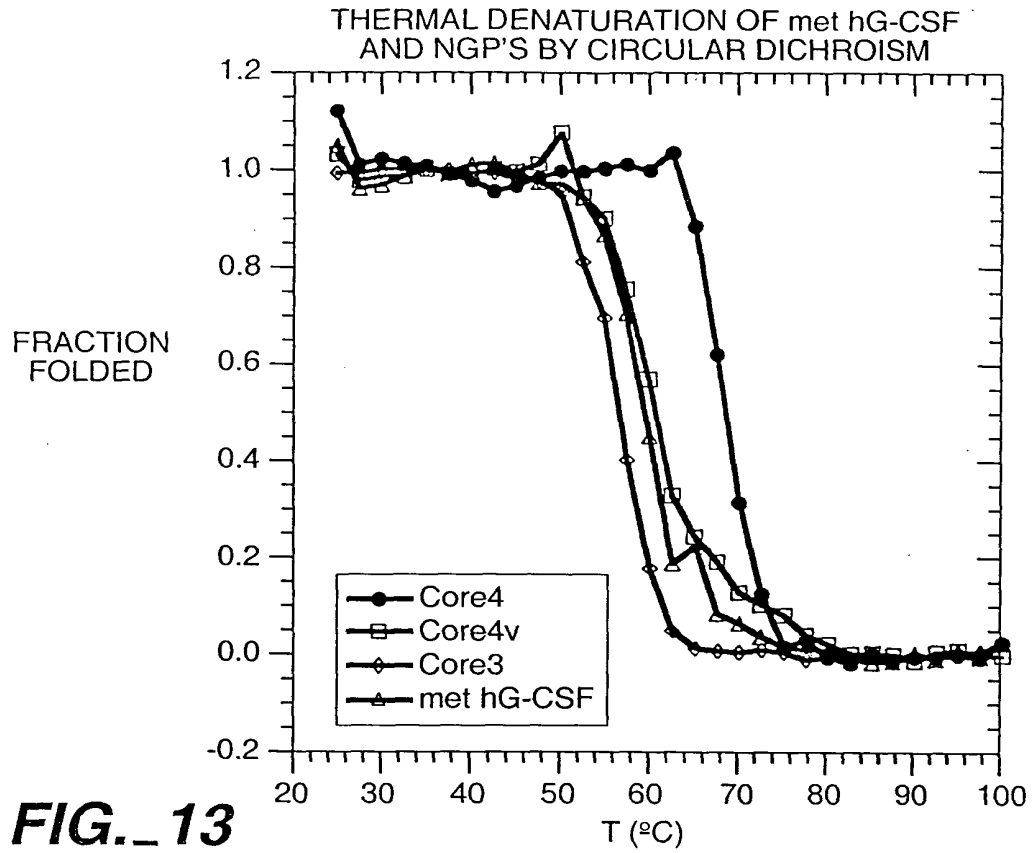
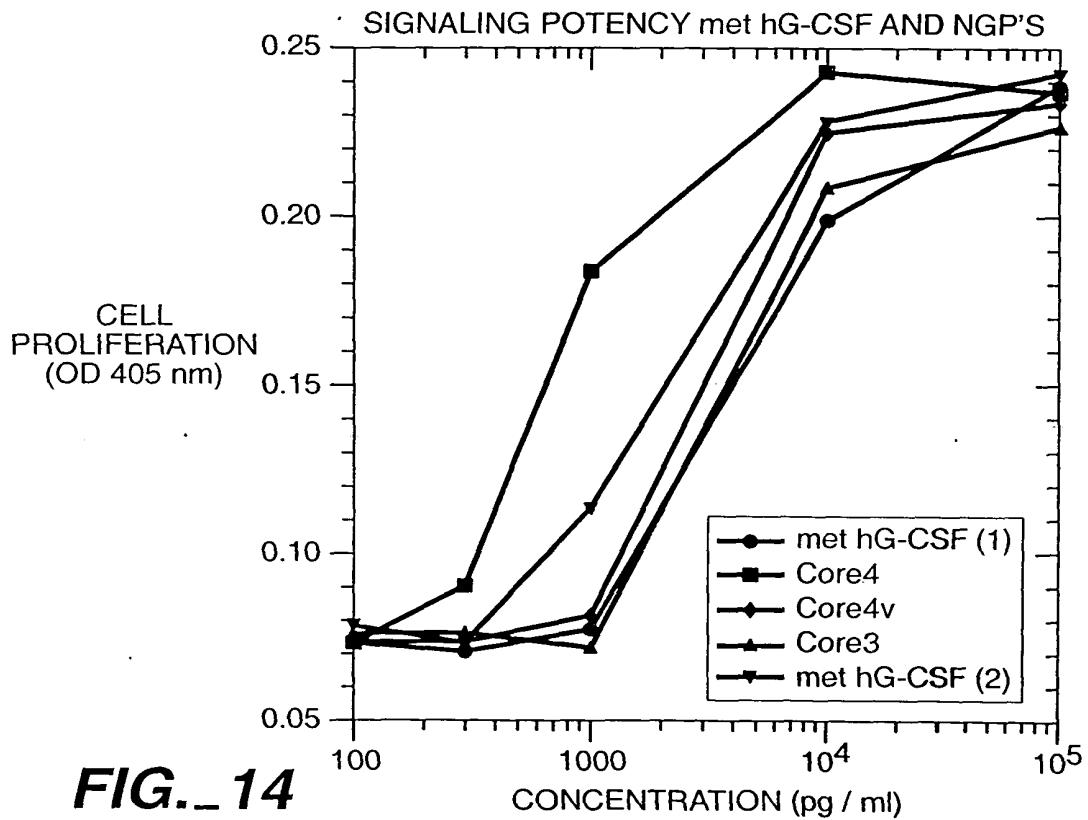
**FIG. 12**-Melting Temperature (T_m)

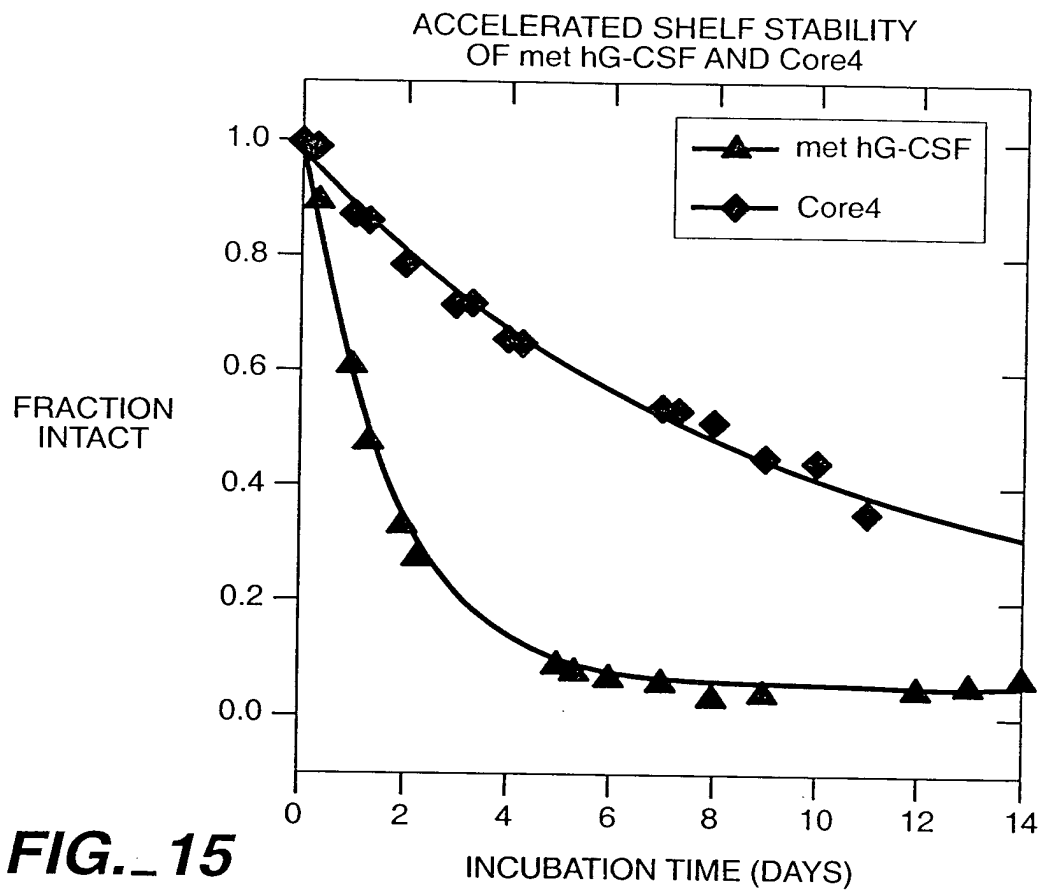
	T_m (°C)	Extinction Coefficient ($M^{-1} \text{ cm}^{-1}$)
hG-CSFwt	60	15720
core4	72	14230
core4v	61	19730
core3	58	14230
sm0*	63	15720
fm4*	63	15720
fm7*	70	14230

* Derived by reverting some core4 or core3
mutant positions to wild type

FIG. 16

15 / 16

**FIG. 13****FIG. 14**

**FIG. 15**